IMAGE PROCESSING METHODS FOR THE MACROSCOPIC ACQUISITION OF HIGH-QUALITY IMAGES OF SURFACES AND TOOLS

F. PUENTE LEÓN

Institut für Meß- und Regelungstechnik, Universität Karlsruhe (TH), Karlsruhe, Germany

SUMMARY

Image processing methods are presented allowing an acquisition of high-quality images of surfaces and tools. The paper focusses on three problems that cause optical systems to fail acquiring images of sufficient quality: unsuitable illumination, limited depth of focus, and visibility problems. In all cases, the accompanying limitations are compensated by obtaining series of images and combining them into an improved result by means of appropriate data fusion techniques. The fusion task is formulated by means of an energy function and can be interpreted as a special case of Bayesian data fusion. The fusion result is obtained by minimization of this function. Due to the complexity of the corresponding optimization, an efficient approximation is proposed which enhances the usability of the approach essentially and nevertheless leads to visually satisfying results. The presented methods were originally developed for acquisition of firearm bullet and cartridge case images, but they are also well-suited for the acquisition of high-quality images of any object in the context of automated visual inspection.

1. INTRODUCTION

Recently, image processing technology has enabled to automate many visual inspection tasks in industrial environments [1]. One important task within this area is concerned with automatic assessment of metallic surfaces and tools such as honed surfaces and milling tools [2, 3, 4]. Commonly a stylus gauge is used for inspection purposes to obtain the surface relief. Alternatively, although less common, images can be obtained with a video camera equipped with a magnifying optical system. The latter method does not yield relief data; however, the optical properties of the surface can be utilized to obtain certain features with maximal contrast. Both techniques are compared in table 1. Of course, the data obtained by both methods complement each other in some ways.

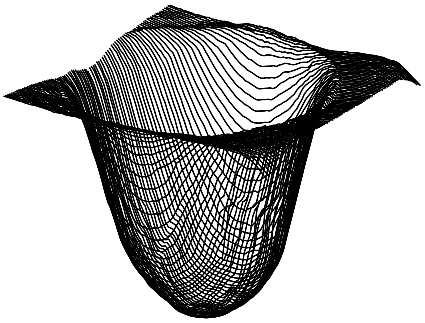
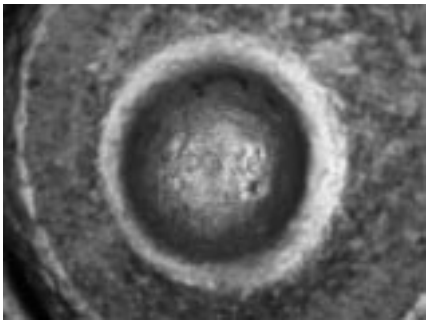
In this paper, image processing methods for the macroscopic acquisition of images of surfaces and tools are presented. To achieve a reliable evaluation of the relevant information about the surface, high-quality images have to be obtained under reproducible conditions. This paper focusses on three problems that cause optical systems to fail acquiring images of sufficient quality for many machine vision applications: unsuitable illumination, limited depth of focus, and visibility problems. In all cases, the accompanying limitations are compensated by acquisition of image series with systematically varying acquisition parameters (illumination angle, object distance, and object pose) and their combination into an improved result by means of appropriate data fusion techniques.

2. IMAGE ACQUISITION PROBLEMS

In automated visual inspection of metallic and spatially structured surfaces, the image acquisition stage bears several difficulties:

- Illumination is a critical aspect regarding image quality and reproducibility of image acquisition conditions. The choice of diffuse illumination generally leads to an undesirable contrast attenuation. Thus, a distant, collimated point source should be preferred. The illumination direction is described by the elevation angle θ and the azimuth φ. Since image intensities highly depend on the illumination direction, optimal lighting would imply that different surface areas are illuminated from different directions, which is extremely difficult.
- The resolution of subtle surface details requires a limited depth of focus. As a result, in

TABLE 1: Comparison between traces from a stylus gauge, and grey level images.

	Stylus gauge	**Grey level image**
		
Depth information	yes	ambiguous
Optical properties	no	yes (e.g. reflectance)
Data acquisition time	high	low
Noise sensitivity	Mechanical stylus: medium Optical stylus: high	low
Reproducibility	high	illumination-dependent
Non-contact measurement	Mechanical stylus: no Optical stylus: yes	yes
Visual interpretation	difficult	possible
Costs	high	medium

many cases, it is not possible to obtain images in which all areas are in-focus.

- The spatial extent of many surfaces leads to distortions due to perspective, and visibility problems. Ideally, images of all surface areas should be acquired under similar geometric conditions.

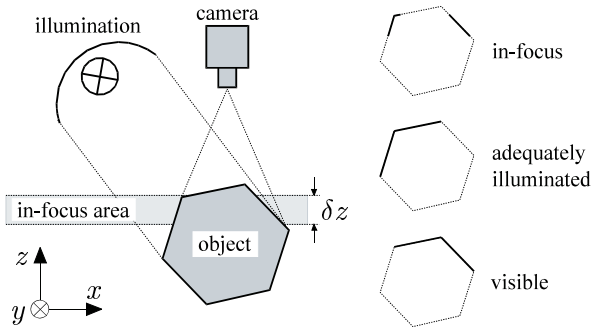


FIG. 1: Image acquisition problems.

The problems arising in the context of image acquisition are illustrated in fig. 1. Quality deficiencies of the image data due to these problems can hardly be compensated in later processing steps. Unfortunately, it is not possible to obtain the whole surface of an object similarly in-focus, and satisfactorily illuminated with only one image. However, image series can be acquired in which the acquisition parameters are varied, so that every surface portion is contained with sufficient quality in one image at least. By means of appropriate data fusion techniques, an image can be obtained in which all areas are contained with sufficient quality.

A series in which only the illumination direction (θ, φ) is varied will be denominated *illumination series*. For a virtual increasement of the depth of focus, a *defocus series* – i.e. a series in which the object distance is varied stepwise – can be acquired. To obtain an image of the whole surface of a spatial object (for example a nearly cylindrical object like a milling tool), a *concatenation* of images acquired by stepwise alteration of the object pose $\bar{\alpha}$ has to be performed. In the next section, efficient techniques for fusing such image series are discussed.

3. IMAGE FUSION

3.1 Image Series Acquisition

The images $d(\bar{x}, \bar{\omega})$ of a series are two-dimensional signals with respect to the location $\bar{x} = (x, y)^{\mathrm{T}}$ indexed with the parameter vector describing the acquisition situation

$$\bar{\omega} = (\varphi, \theta, \zeta, \bar{\alpha}, \ldots)^{\mathrm{T}},$$

where φ and θ represent azimuth and elevation angle of the illumination direction respectively, ζ the object distance, and $\bar{\alpha}$ the object pose. Additional parameters, like the integration time of the video camera or the wavelength of the used light, could also be taken into account, if necessary.

Before an image series can be acquired, it has to be determined how the parameter space has to be sampled. The goal is to obtain all surface areas in good quality with as few images

$$D = \{d(\bar{x}, \bar{\omega}_i), i = 0, \ldots, n\}$$

as possible. This problem highly depends on the object geometry as well as the surface texture and cannot be dealt with in detail here. Some aspects of this problem are discussed in [5, 6, 7].

An answer can be given for defocus series. Every portion of the surface $z(\bar{x})$ will be contained at least once in focus, if the following conditions hold:

$$\zeta_i = \zeta_0 + i\Delta z, \quad i = 0, \ldots, n,$$

$$\Delta z \leq \delta z, \quad \zeta_0 \leq z(\bar{x}) \leq \zeta_n, \tag{1}$$

where δz denotes the depth of focus.

For illumination series of textured surfaces consisting of a band of straight, parallel grooves, it is not necessary to sample the illumination space two-dimensionally, because such surfaces only show a high contrast if they are illuminated perpendicularly. Thus, only the elevation angle θ has to be varied [8]. However, in most cases both illumination angles will have to be varied.

3.2 Fusion Approach

The literature on data fusion is extensive, indicating the interest in this topic. However, many of the approaches are *ad hoc*. In this work, a systematic data fusion approach is used which is based on the minimization of a so-called "energy function" [9]

$$E = E_{\mathrm{D}}(D, r) + \lambda E_{\mathrm{C}}(r), \quad \lambda > 0. \tag{2}$$

$E_{\mathrm{D}}(D, r)$ models the relationship between the given image data (i.e. the image series)

$$D = \{d(\bar{x}, \bar{\omega}_i), i = 0, \ldots, n\} \tag{3}$$

and the fusion result $r(\bar{x})$. $E_{\mathrm{C}}(r)$ models desired characteristics of the fusion result $r(\bar{x})$ or those known a priori. The regularization parameter λ serves to weight both components.

The energy terms $E_{\mathrm{D}}(D, r)$ and $E_{\mathrm{C}}(r)$ are to be defined in such a way that the result is more desirable, the lower the energy is. Consequently, E has to be minimized to obtain $r(\bar{x})$.

By defining Gibb's densities, a connection of this approach with the Bayesian fusion theory and the Markov Random Fields theory can be achieved [9]. Thus, methods for solving inverse problems of statistical mechanics can be also utilized.

3.3 Fusion of 1-D Illumination Series

For a better understanding, a fusion algorithm for illumination series will be presented first. Following, the fusion task will be generalized in such a way that it will fit within the framework discussed in the last section. It will be shown that the method represents a very efficient approximation for the solution of the fusion problem by energy minimization.

In this section one-dimensional series – i.e. series in which only one parameter is varied – will be treated. An extension to the two-dimensional case will be given in the next section. Fig. 2 shows the structure of the fusion algorithm for the case of a varying azimuth φ.

The principle of the fusion algorithm consists in the selection of the best illuminated image segments of the series for each location based on a local criterion C. The local grey-level variance and the local entropy are suitable criteria C, if a high-contrast fusion result $r(\bar{x})$ is desired. The selected illumination direction, which is stored for each location $\bar{x}$ in the so-called *illumination map*

$$\tilde{\varphi}(\bar{x}) = \arg\max_{\varphi_i} C\{d(\bar{x}, \varphi_i)\}, \tag{4}$$

has to be a spatial function varying slowly compared to the signal of interest. This is necessary to avoid artifacts in the fusion result. To assure that this condition is satisfied, a smoothing of the illumination map with a binomial low-pass filter is performed [10]:

$$\phi(\bar{x}) = \angle \mathrm{LP}\{\mathrm{e}^{j\tilde{\varphi}(\bar{x})}\}. \tag{5}$$

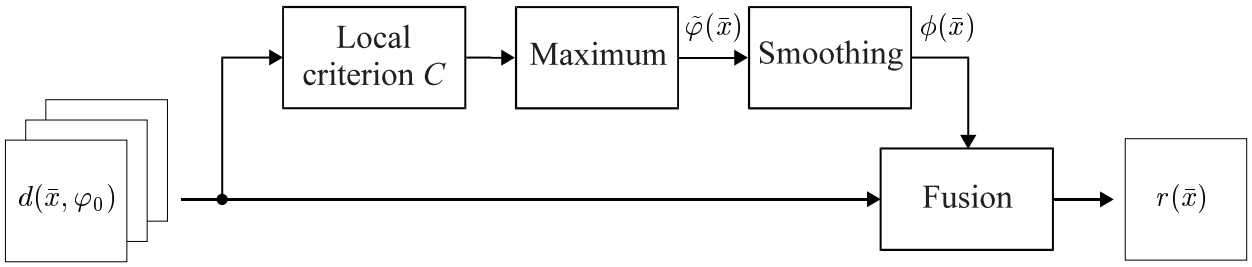


FIG. 2: Structure of the algorithm for fusing 1-D illumination series.

In this step, the cyclicity of φ has to be taken into account, because $\varphi = \varphi + 2\pi k,\ k \in \mathbb{Z}$ holds. Thus, not $\tilde{\varphi}(\bar{x})$ itself, but the complex pointer $\exp(j\tilde{\varphi}(\bar{x}))$ has to be smoothed. The resulting function $\phi(\bar{x})$, which denotes the best-suited local illumination direction, is the angle of the complex result [11].

The actual fusion is performed by weighted superposition of two adjacent images $d(\bar{x}, \varphi_i)$ with a linear interpolator γ taking the best local illumination direction $\phi(\bar{x})$ into account:

$$\begin{aligned} r(\bar{x}) &= \sum_i d(\bar{x}, \varphi_i)\gamma(\phi(\bar{x}) - \varphi_i) \\ &= \frac{\phi(\bar{x}) - \varphi_l}{\varphi_{l+1} - \varphi_l} d(\bar{x}, \varphi_l) + \\ &\quad \frac{\varphi_{l+1} - \phi(\bar{x})}{\varphi_{l+1} - \varphi_l} d(\bar{x}, \varphi_{l+1}). \end{aligned} \tag{6}$$

The interpolation takes care of a smooth transition between φ-neighbouring images; see fig. 3. The narrow extent of γ provides for an averaging of only similarly illuminated images. Thus, an undesirable contrast loss due to destructive interferences of light and shadow in different images of the series is avoided.

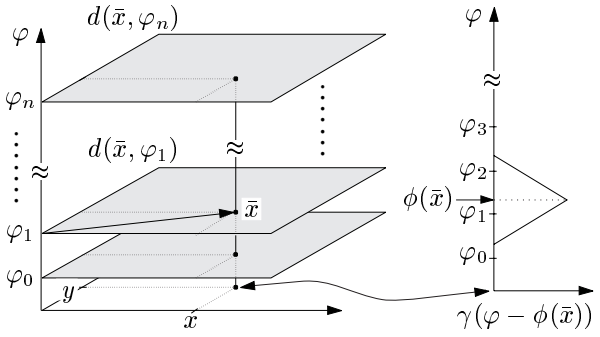


FIG. 3: Selection of the best illuminated image segments for fusion.

Three properties of the proposed fusion method are responsible of its good performance:

1. for each location $\bar{x}$, the fusion result $r(\bar{x})$ resembles the best illuminated image $d(\bar{x}, \varphi_i)$ of the series;
2. the smoothness of the selected illumination direction $\phi(\bar{x})$ guarantees that no artifacts are contained in the resulting image $r(\bar{x})$;
3. the resulting image achieves globally good results in the sense of the local measure C.

By formulating energy terms that penalize the nonfulfillment of any of these conditions, an energy function of the form of eq. (2) can be obtained:

$$\begin{aligned} E &= \sum_i \sum_{\bar{x}} (r(\bar{x}) - d(\bar{x}, \varphi_i))^2 \gamma(\varphi_i - \phi(\bar{x})) \\ &\quad + \lambda_1 \sum_{\bar{x}} (\mathrm{HP}\{\phi(\bar{x})\})^2 \\ &\quad - \lambda_2 \sum_{\bar{x}} C\{r(\bar{x})\} \\ &= E_\mathrm{D}(D, r, \phi) + \lambda_1 E_\mathrm{S}(\phi) \\ &\quad + \lambda_2 E_\mathrm{C}(r). \end{aligned} \tag{7}$$

This equation represents a compact, implicit formulation of the fusion problem, in which all known and desirable characteristics of the magnitudes involved in the fusion process as well as their mutual relations are given.

The first addend $E_\mathrm{D}(D, r, \phi)$ in eq. (7) provides for data proximity to $r(\bar{x})$. To fulfill the smoothness constraint for the optimal illumination angle $\phi(\bar{x})$, the second addend $E_\mathrm{S}(\phi)$ penalizes high energy components of $\phi(\bar{x})$ by measuring the energy of the high-pass filtered signal $\mathrm{HP}\{\phi(\bar{x})\}$. Consequently, this addend represents a smoothness constraint for the optimal illumination angle $\phi(\bar{x})$. The third addend $E_\mathrm{C}(r)$ checks whether the local criterion C leads to high values in the fusion result $r(\bar{x})$ globally.

If we have a look at the algorithm shown in fig. 2 again, the analogy between its blocks and the addends of eq. (7) becomes more clear. In both left

upper blocks basically $E_{\mathrm{C}}(r)$ is minimized. The smoothing performed in the right upper block above all provides for a minimization of $E_{\mathrm{S}}(\phi)$. Subsequently, the fusion block takes care of a smooth transition between the best images of the series, being equivalent to a minimization of $E_{\mathrm{D}}(D, r, \phi)$.

For the assumptions made here, the minimization of E with respect to r and ϕ would lead to the optimal fusion result at the expense of a very high computation time. The fusion strategy proposed instead, however, represents an *efficient approximation* of the energy minimization approach based on a separate optimization of the addends of eq. (7) and with no need to consider the weighting factors λ_i.

3.4 Fusion of 2-D Illumination Series

In the more general case of varying the azimuth φ as well as the elevation angle θ, the fusion approach given in eq. (7) has to be extended accordingly:

$$\begin{aligned} E &= E_{\mathrm{D}}(D, r, \phi, \vartheta) + \lambda_1 E_{\mathrm{S}}(\phi) \\ &\quad + \lambda_2 E_{\mathrm{C}}(r) + \lambda_3 E_{\mathrm{S}}(\vartheta). \end{aligned} \tag{8}$$

An additional term proportional to $E_{\mathrm{S}}(\vartheta)$ provides for smoothness of the locally optimal elevation angle $\vartheta(\bar{x})$. Furthermore, the linear interpolator γ contained in the first term $E_{\mathrm{D}}(D, r, \phi, \vartheta)$ has to be expanded to the two-dimensional case. The narrow extent of $\gamma(\phi, \vartheta, \varphi_i, \theta_i)$ only allows images in a neighbourhood around the locally optimal illumination direction $(\phi(\bar{x}), \vartheta(\bar{x}))$ to contribute to the fusion result $r(\bar{x})$.

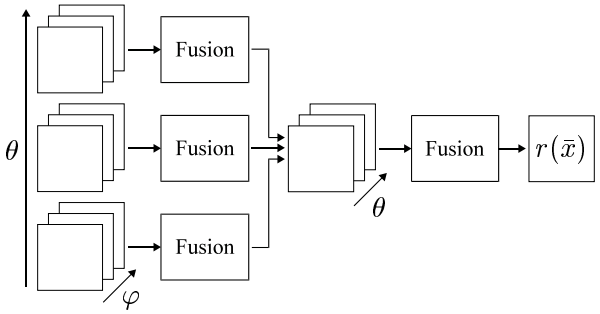


FIG. 4: Fusion of 2-D illumination series.

If the illumination space is sampled properly, the fusion task can be treated one-dimensionally, and consequently simplified; see fig. 4. For certain elevation angles θ_i, $i = 0, \ldots, n$ illumination series are acquired by varying the azimuth φ. In the first fusion stage, all image series have to be fused with respect to φ. Each of the fusion blocks corresponds to the whole fusion algorithm shown in fig. 2. The resulting images represent a new image series. By fusing this series with respect to θ, the fusion result $r(x)$ is obtained.

3.5 Fusion of Defocus Series

The fusion of defocus series can be performed similarly to 1-D illumination series, if the optimal elevation angle $\phi(\bar{x})$ is replaced by the optimal object distance $\xi(\bar{x})$, and the angles φ_i by the actual object distances ζ_i:

$$E = E_{\mathrm{D}}(D, r, \xi) + \lambda_1 E_{\mathrm{S}}(\xi) + \lambda_2 E_{\mathrm{C}}(r). \tag{9}$$

The qualitative meaning of the energy terms as well as the local criterion C remain the same as with illumination series. In particular, a certain smoothness of the optimal object distance $\xi(\bar{x})$ has also to be postulated here, because for the empirical estimation of focus by means of the criterion C, spatial averaging in a neighbourhood is required. Thus, within this neighbourhood, a nearly constant object profile has to be assumed.

The task of fusing defocus series can be also performed efficiently with the suboptimal strategy proposed in section 3.3. In this case, instead of the illumination map, a so-called *depth map* $\tilde{\zeta}(\bar{x})$ is obtained, which represents a measure of the surface profile.

An additional advantage of fusing defocus series is that, since the distance to the selected surface areas is nearly constant in case of a small depth of focus δz, the image acquisition process has telecentric properties.

3.6 Image Concatenation

In many cases, the geometry of an object will complicate acquiring its surface with only one image. Moreover, it could be impossible to obtain an image of a large surface with the required resolution. In such cases, multiple images have to be acquired by varying the object pose $\bar{\alpha}$, so that the same border area is always contained in two consecutive images; see fig. 5.

In the present case, an image of the whole circumferential surface of a nearly cylindrical object like a milling tool has to be obtained. Consequently, in

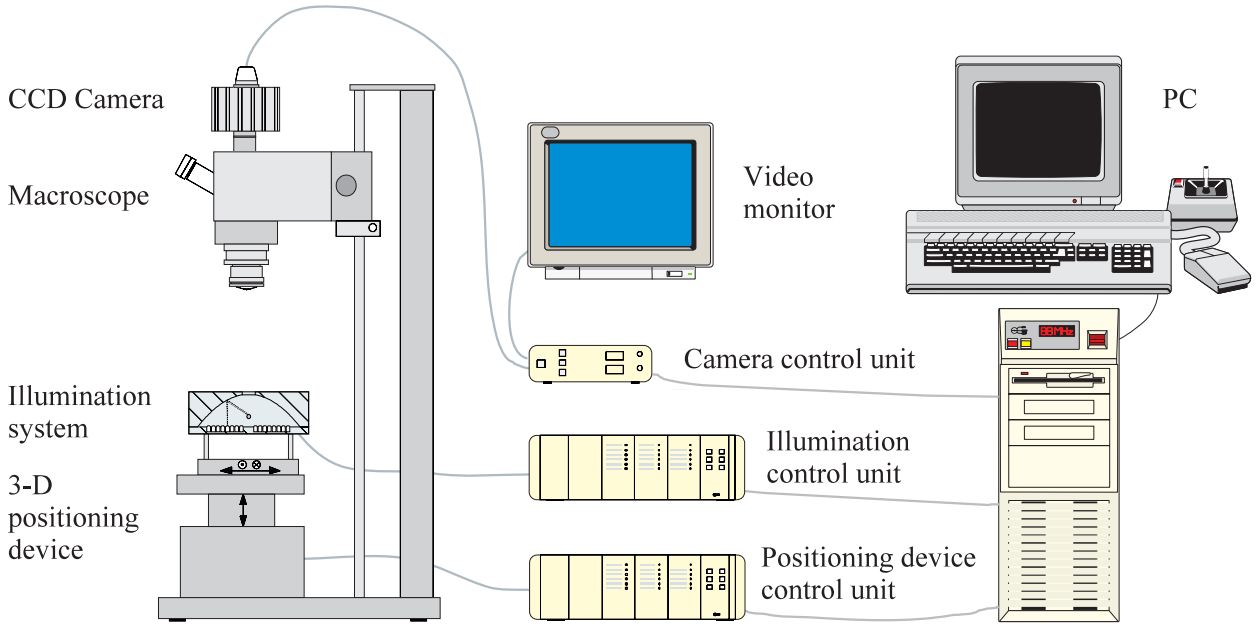


FIG. 6: Experimental set-up.

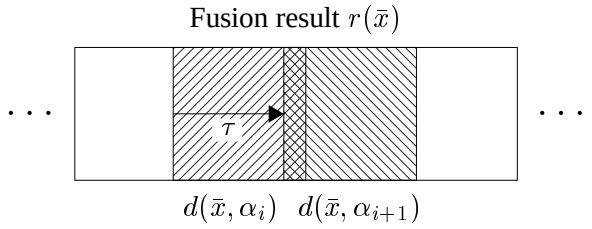


FIG. 5: Concatenation principle.

the image series only the rotational position α of the object needs to be varied:

$$\begin{aligned} D &= \{d(\bar{x}, \alpha_i), i = 0, \ldots, n\}, \\ \alpha_i &= \alpha_0 + i\Delta\alpha. \end{aligned} \tag{10}$$

Due to the knowledge of the variation of α, the translations τ between consecutive images are also approximately known. However, for a precise reconstruction of the surface, τ has to be determined more exactly by means of cross-correlation of the overlapping stripes [12].

To guarantee similar illumination and geometric conditions in the overlapping areas, $\Delta\alpha$ has to be chosen small enough. In the overlapping areas, a weighted averaging between consecutive images is performed to suppress small grey-level fluctuations.

4. Experimental Results

In this section, experimental results of the discussed fusion methods are presented and compared with images which can be obtained without data fusion. To obtain image series, an automated system was set up which consists of a flexible illumination module, a commercial macroscope, and a 3-D positioning device; see fig. 6. The illumination system is composed of a platform in which 256 LEDs are placed, and a parabolic reflector in the focus of which the object is fixed; see fig. 7 [13]. The location of an LED on the platform determines the direction (θ, φ) from which the object surface is illuminated. By variation of these parameters, any area on the object surface can be acquired with maximum contrast. An opening in the reflector allows image acquisition with a macroscope and a CCD camera. All images throughout this paper were digitized with 512×512 pixels, and 8 bit grey levels.

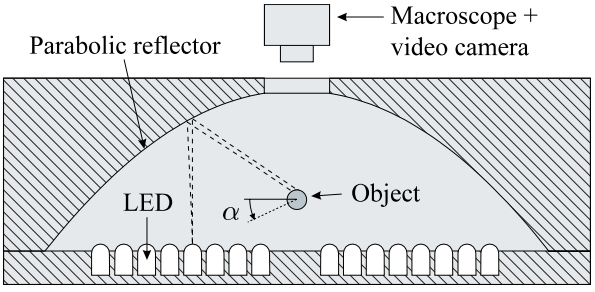


FIG. 7: Illumination system.

In fig. 8a, an image of an illumination series consisting of 20 images ($\Delta\varphi \approx 5.6°$) of an end-milling texture can be seen. Due to illumination, the texture shows a bright stripe [8]. To fuse the image

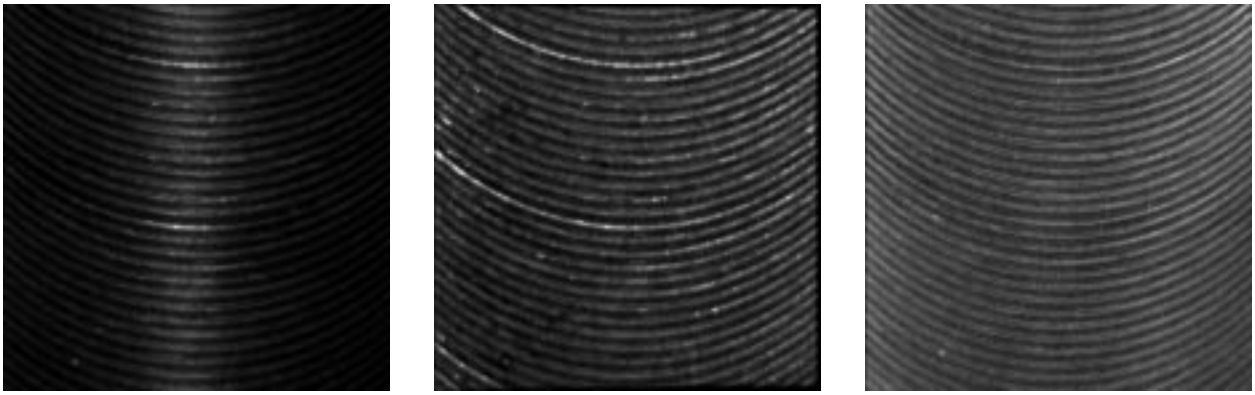


FIG. 8: End-milling surface: a) image no. 8 of the illumination series; b) fusion result (criterion C: variance in a 5×5-mask; smoothing of $\tilde{\varphi}(\bar{x})$ with a binomial filter of size 49×49); c) diffuse lighting.

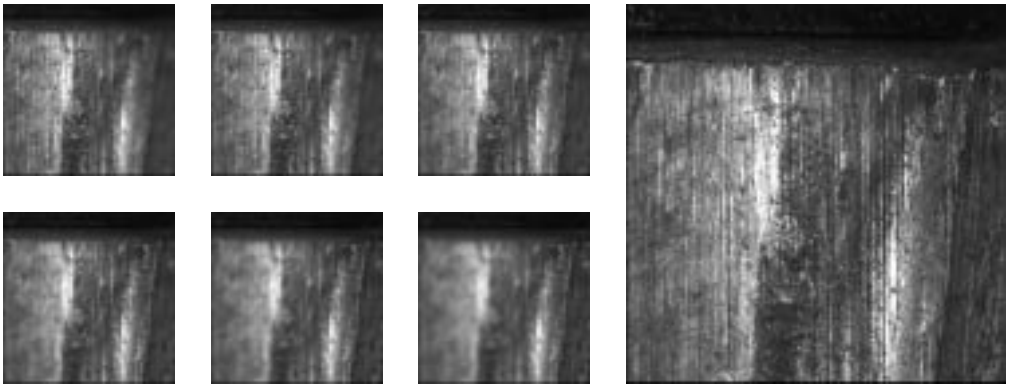


FIG. 9: Firearm bullet: l.) defocus series (images 2, 5, 8, 11, 14, and 17); r.) fusion result (criterion C: variance in a 5×5-mask; smoothing of $\tilde{\zeta}(\bar{x})$ with a binomial filter of size 21×21).

FIG. 10: Milling tool.

series, C was chosen as the local grey-level variance in a mask of size 5×5. The illumination map $\tilde{\varphi}(\bar{x})$ was smoothed with a circular binomial filter with an impulse response of 49×49 pixels. In the fusion result fig. 8b, the whole surface is illuminated much better, hitherto hidden details become visible, and the stripe-like inhomogeneity can no longer be recognized. By comparison of the fusion result with the same diffusely illuminated surface (fig. 8c), it can be stated that in the fusion result the grooves are contained with much higher contrast. Thus, the surface quality can be assessed more accurately.

In the left side of fig. 9, six images of a defocus series consisting of 20 images ($\Delta z = 78.5\mu\mathrm{m}$) of a firearm bullet are shown. All images of the series were obtained with diffuse illumination. The same criterium C as in fig. 8 was chosen in this case. The depth map $\tilde{\zeta}(\bar{x})$ was smoothed with a binomial filter with an impulse response of 21×21 pixels. Obviously, in the fusion result, which is depicted in fig. 9c, all surface areas are in-focus, in contrast to the single images of the series. By means of fusion, a virtual increase of the depth of focus could be achieved.

In fig. 11, the concatenation result of an image series of the milling tool represented in fig. 10 is shown. In the corresponding series, the object distance ζ ($\Delta z = 210.5\mu\mathrm{m}$), and the rotational position α of the object ($\Delta\alpha = 1.8°$) were varied. The signals $\tilde{\zeta}(\bar{x})$ were smoothed with a binomial filter of size 21×21. The other parameters were chosen as in fig. 9. Note that even deep areas of the tool are contained sharply focussed. Particularly, in the right third of the image, a breakage of a cutting edge can be clearly seen.

5. Summary and Conclusions

In this paper, methods for obtaining high-quality images of surfaces and tools have been presented. To compensate the limitations of optical systems, image series were obtained by varying the acquisition parameters systematically. By means of data

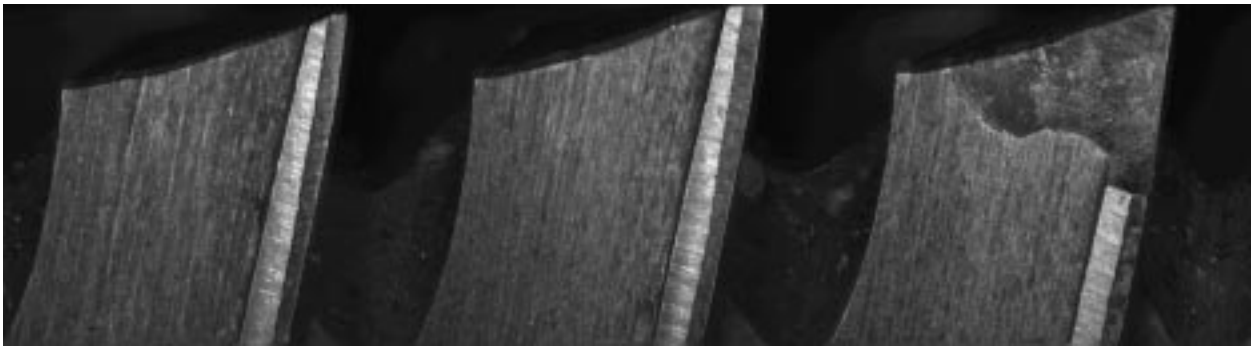


FIG. 11: Milling tool: concatenation result.

fusion, images were combined to an improved result which could not have been acquired physically with only one image. The fusion task has been formulated by means of an energy function. On the one hand its addends describe all known and desirable relations between the image series $\{d(\bar{x}, \omega_i)\}$ and the fusion result $r(\bar{x})$, and on the other hand the known and desirable properties of $r(\bar{x})$ and other relevant magnitudes involved in the fusion process. By minimization of this function, the optimal fusion result with respect to the assumptions met was obtained. In our case, the structure of the energy function allowed to perform the computationally expensive optimization by means of an efficient approximation.

The performance of the proposed algorithms has been demonstrated with images of metallic surfaces and tools. However, the methods presented are also suitable for acquisition of high-quality images of any other object for automated visual inspection purposes. As a rule, surface features could be obtained much more robustly and with higher contrast by means of the fusion methods presented. Hence, the increased effort in image acquisition appears to be absolutely reasonable in many computer vision tasks, where high-quality images are needed.

REFERENCES

[1] T. S. NEWMAN, A. K. JAIN, A Survey of Automated Visual Inspection, *Computer Vision and Image Understanding* **61** (2), pp. 231–262 (1995).

[2] J. BEYERER, Detektion von Fräserschaden durch Texturanalyse der Werkstückoberfläche, *Technisches Messen* **60**, pp. 419–424 (1993).

[3] J. BEYERER, *Analyse von Riefentexturen*, VDI-Verlag, Düsseldorf, 1994.

[4] J. BEYERER, F. PUENTE LEÓN, Detection of defects in groove textures of honed surfaces, *International Journal of Machine Tools & Manufacture* Vol. 37, No. 3, pp. 371–389, (1997).

[5] R. MALZ, *Codierte Lichtstrukturen für 3-D-Meßtechnik und Inspektion*, Dissertation, Universität Stuttgart, 1992.

[6] G. H. TARBOX, S. N. GOTTSCHLICH, Planning for Complete Sensor Coverage in Inspection. *Computer Vision and Image Understanding* **61** (1), 84–111 (1995).

[7] S. YI, R. M. HARALICK, L. G. SHAPIRO, Optimal Sensor and Light Source Positioning for Machine Vision. *Computer Vision and Image Understanding* **61** (1), 122–137 (1995).

[8] J. BEYERER, F. PUENTE LEÓN, Suppression of inhomogeneities in images of textured surfaces, *Optical Engineering* **36** (1), pp. 85–93 (1997).

[9] J. J. CLARK, A. L. YUILLE, *Data Fusion for Sensory Information Processing Systems*, Kluwer, Boston, 1990.

[10] B. JÄHNE, *Digitale Bildverarbeitung*, Springer, Berlin, 1993.

[11] K. V. MARDIA, *Statistics of Directional Data*, Academic Press, London, 1972.

[12] H. P. BÄHR, T. VÖGTLE, *Digitale Bildverarbeitung*, Wichmann, Karlsruhe, 1991.

[13] R. MALZ, *Verfahren zum beleuchtungsdynamischen Erkennen und Klassifizieren von Oberflächenmerkmalen und -defekten und Vorrichtung hierzu*, Offenlegungsschrift DE 41 23 916 A1, Deutsches Patentamt, 1992.

Contact point: Dipl.-Ing. Fernando Puente León, Institut für Meß- und Regelungstechnik, Universität Karlsruhe (TH), Postfach 6980, 76128 Karlsruhe, Germany, Phone Int +49 721 608-3604, Fax Int +49 721 661874, E-mail: f.puente@ieee.org